AD NUMBER

AD297375

NEW LIMITATION CHANGE

TO

Approved for public release, distribution unlimited

FROM

Distribution authorized to U.S. Gov't. agencies only; Administrative/Operational Use; Nov 1962. Other requests shall be referred to the Department of the Army, Attn: Public Affairs Office, Washington, DC 20310.

AUTHORITY

CRD D/A ltr, 25 Feb 1963

NOTICE: When government or other drawings, specifications or other data are used for any purpose other than in connection with a definitely related government procurement operation, the U. S. Government thereby incurs no responsibility, nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use or sell any patented invention that may in any way be related thereto.

TECHNICAL PAPER
RAC-P-74
NOVEMBER 1962

RESEARCH ANALYSIS CORPORATION

Lanchester as Force in History:

An Analysis of Land Battles of the Years 1618—1905

by
D. Willard

PREPARED FOR THE
DEPARTMENT OF THE ARMY
Contract No. DA 44-188-ARO-1

Copy ______ of 250

ECONOMICS AND COSTING DIVISION
TECHNICAL PAPER RAC-TP-74
Published November 1962

Lanchester as Force in History:

An Analysis of Land Battles of the Years 1618—1905

by
D. Willard

RESEARCH ANALYSIS CORPORATION
6935 Arlington Road, Bethesda, Md., Washington 14, D.C.

Received for Publication
2 October 1962
Published
November 1962
by
RESEARCH ANALYSIS CORPORATION
6935 Arlington Road
Bethesda, Md.

FOREWORD

The work reported here was undertaken as a part of RAC Project 91.1.5: "Limited War Capability Study – 1967." The analysis had not been sufficiently advanced at the termination of that project to allow the findings to be considered for inclusion in the overall report.

The decision to undertake this analysis was made by Dr. Wilbur B. Payne, coordinator of the study, who saw its usefulness as guidance for certain techniques of analyzing or simulating the results of war games. Dr. Payne followed this work closely and provided considerable aid in establishing original guidelines, in interpreting preliminary results, and in developing the conclusions. Dr. Philip H. Lowry has provided valuable assistance throughout the study. His long-standing interest in analysis of the tables in Bodart's Lexikon was an important part of the initial stimulus. The author has benefited materially from discussions with other Yorkists and Lancastrians on the RAC staff, particularly Dr. Hugh M. Cole and Dr. George S. Pettee. In the end, however, the responsibility for errors of analytical procedures, of data, and of findings lies solely with the author.

The author is grateful to Mr. George E. Clark Jr. and to the staff of the Computation Laboratory for expeditiously transferring a mass of data in Bodart to punched cards, and for assistance in programming the analytical routines for execution. He is obliged to the staff of the Strategy and Tactics Analysis Group, and particularly to Mr. Daniel Belsole of the Computation Division, for most thoroughgoing cooperation during the trial and execution of the program on their IBM 7090. Finally, he thanks Dr. Irving H. Siegel who tolerated this diversion in his Division.

CONTENTS

TABLES

PROBLEM

Without prejudice to the hypothesis that either of Lanchester's laws—linear or square—accurately depicts the attrition of forces over short periods of time in the engagements of small land-combat units (e.g. platoons or companies), it may be asked whether an elementary form of these laws, or a simple generalization of them, adequately describes the course and outcome of battles between armies. Specifically, the goal of the investigation reported here was to determine, by an examination of historical military data, the extent to which Lanchester's equations are an expression of a general property of battle.

BACKGROUND

In order to simulate combat (as with a war game) or to interpret the results of combat real or simulated it is desirable to have a set of critieria that will help to answer the question: "Which side will win this battle?" The criteria should be objective; i.e., expressed in terms of observables such as the number of casualties caused, prisoners captured, territory controlled, or other tangible evidence of a victorious posture. Simplicity of application of the criteria also argues that they should be quantitative, and among the simplest measures of attrition of forces are those proposed by F. W. Lanchester[1] in the form of the equations that bear his name.

In attempting to apply Lanchester's equations to the outcome of war games it was found necessary to invoke some principle—historical precedent was chosen—in order to fix the values of certain quantities in a generalized form of the equations. Weiss[2] has carried out the only comparable analysis; for a small sample of recent battles he presents evidence of the relevance of Lanchester's square law On the other hand he questions—and so does Snow[3]—the validity of linear or square law when applied to arbitrary land battles. Helmbold[4] has derived some of the parameters of 92 battles from an extended period,* but he assumes that the battles analyzed necessarily belong to a particular Lanchester species (again the square law). The major surveys of the statistics of armed combat (up to or including WWII) have been largely limited to the casus belli and not to the circumstances that led to the termination of combat.[5,6]

It was not possible in the time available to expiscate from their diverse sources the appropriate details of modern battles; on the other hand, Bodart[7]

*The most significant part of his work did not come to our attention until this paper was in proof. It appears that the conclusions of Ref 4 and of this study are complementary.

has provided a comprehensive statistical survey of 300 years of warfare (through the Russo-Japanese War). The results reported here have been developed from his tables. Smith and Donovan[8] made an examination of the same source for the casualty ratios and the effect of force ratio on chances of victory; otherwise the source examined here appears to be a fallow field and the treatment given it in this paper, unique.

DISCUSSION

For a quick reading of this paper the reader is directed to the "Introduction" and to Secs 1 and 4. Here he will find a description of the data extracted from Bodart, the methods of grouping the battles for analysis, and four tables (Tables 2, 3, 6, and 8) that could be developed without regard to any one theory of battle. The battles were assigned to either of two categories: I, meeting engagements and similar combat; II, sieges, attacks on fortified positions, and similar combat. No explicit designation of either combatant as attacker was made, although for the battles of Category II the attacker was generally the stronger.

Table 2 shows that Category I combatants joined battle most frequently when the force ratio was nearly 1 to 1, rarely at disparate strengths. According to Table 3, victory in a Category I battle was almost independent of force ratio. For Category II the reverse was true; force ratio had little to do with the joining of battle but rather more with the outcome of the battle.

Table 8 shows that for the 84 most recent battles of Category I force ratio has no effect on outcome until it exceeds 4 to 1.

Though it has nothing to do with force ratios, Table 6 shows how frequently bloody battles have occurred. To enter the table, one must first compute the ratio of killed and wounded to initial strength for each side and select the larger number. In the table he will find the fraction of the battles studied for which this number is less than selected values.

The results that deal with force ratio have an important bearing on the rest of the paper: if force ratio has little effect on outcome of battle, Lanchester's equations cannot be expected to be well validated.

Forget this particular warning temporarily. It is still not at all obvious that either the linear law or the square law is correct as a rule. A first step in the direction of generalization is the construction of a form of the equations that contains the two laws as special cases. From initial- and final-strength data for battles to which this general form does apply, it is possible to determine the value of a parameter of the equations—here called γ—which distinguishes among linear, square, and "intermediate" laws. Methods for finding the best value of γ and for finding how well the theory supports the data are described in Secs 2 and 3; the results of the tests are reported and discussed in Secs 5 and 6.

Findings

(a) Results are insensitive to the choice of γ, but if it has a single best value this value seems to imply that the rate at which one side suffers casualties increases as its opponent's strength is reduced. This description of an effect is not to be taken as an explanation of it. The observation holds for each Category when it is considered intact and also when it is split into a few sections chosen according to the year of the battle, the combined strength of opponents, the force ratio, or the ratio of casualties to initial strength. (For an exception, see b.) The effect is more pronounced for battles of Category II.

(b) Those battles in which the exchange ratio is most nearly unity (as computed from linear or square law) also show the nearest fit to these laws. An interesting by-product of the analysis of Sec 5 is the table (App B) of those battles in Category I which most nearly fit these laws. With the exception of Blenheim and Marengo they seem historically insignificant.

(c) Some evidence suggests that the rate of attrition of the winner's forces depends on the loser's strength in a manner different from the dependence of the loser's casualty rate on the winner's strength. An extreme example of this effect would be that battle in which the loser presents himself as a target according to the linear law while the winner appears as a square-law target. Of course the statistical tests reported here support no such clear-cut differentiation, nor is there any simple interpretation of this result.

(d) The findings in a and c are insensitive to errors of 10 percent in the casualty data, as demonstrated by altering these numbers either systematically or randomly by such amounts.

(e) A stochastic version of Lanchester's equations correctly "predicts" the outcome of 8 battles out of 10, where sheer chance would give correct results for 5 battles. It was possible to achieve this accuracy by using a value of the exchange ratio computed from final and initial strengths.

(f) As a result of e the stochastic theory can be "asked" to predict the contents of Table 3. The results are shown in Table 14. It will be seen that the battles of Category II are rather well explained, but those of Category I are less successfully explained because theory predicts a dependence on force ratio. The results of e and f are entirely unaffected by a choice of γ.

(g) The exchange ratio has proved a mercurial concept, leading to several unfruitful analyses at the early stages of this study. A method of computing it for a single battle served as a control in two tests (b and e), but attempts to isolate a best value for a Category or a subcategory were unsuccessful.

CONCLUSIONS

1. Since the battles of Category II show a higher correlation with theory than do those of Category I, the properties that differentiate these groups may provide a key to better understanding of attrition of forces in land warfare.

SUMMARY

One of these properties may be posture (offensive or defensive), but alone it is insufficient to achieve accord between facts and theory.

2. In general, force ratio has had little to do with determining the outcome of the battles studied.

3. Lanchester's square law is the poorest among poor alternative choices of deterministic laws, though results are insensitive to γ.

4. By elimination it is presumably the exchange ratio E that has controlled the outcome of these battles and is in some fashion responsible for the extent to which a stochastic form of the theory explains Table 3. Although Lanchester offers an expression that apparently yields by hindsight a satisfactory estimate of E there is no theory for E. The writer concludes that in the absence of any method of predicting E reliably there is little value in a simple version of Lanchester's equations as a predictive tool where the only known quantities are initial strengths.

Since these conclusions are mainly negative in character they are really questions to be asked about the outcome of more recent battles. Considering that the largest collection of battles yet analyzed yields the lamest defense of Lanchester, it may be asked if these results do not simply furnish an additional proof that large quantities of mediocre data are no substitute for a limited amount of accurate data? Maybe, but there is strong reason to believe that the testing of contemporary battles against these findings would be inconclusive if it were to rely on a small quantity of data, whatever the quality.

As a photograph can be misleading because the third dimension has been suppressed, so presumably the peculiar results described above are peculiar because the theory lacks an essential dimension—a hypothesis. It is unlikely that poor data are entirely responsible, though more and better data can shed the light that converts a two-dimensional illusion into a meaningful three-dimensional reality.

At the present time no satisfactory explanation of the deficiencies of Lanchester's equations is known. The following conjecture is unsupported but open to testing: Although human behavior patterns (constraints) exist for which there are statistical laws, the classic situations to which the linear or the square law may be expected to apply (Weiss[2]) represent battles in which the functions of command and control seem to be missing. The equations seem to be averaging over dehumanized events. If this is so, then a probabilistic theory, if it predicts that the most likely state of affairs is conformity to Lanchester's equations, may be correct only in the absence of command or of opportunity for effective control. The presence of these factors may make conformity least likely. *

5. Without prejudice to the status of Lanchester's equations as they refer to small units in combat, it is felt that justification for their use in large-scale situations has not been demonstrated.

*For a different way of expressing concern on this topic, see Cline,[9] Sec. 4.

LANCHESTER AS FORCE IN HISTORY: AN ANALYSIS OF LAND BATTLES OF THE YEARS 1618–1905

GLOSSARY OF SYMBOLS

A_s	actual number of battles won by the stronger side
E	(differential) exchange ratio
f_m	m_c/m_0
f_n	n_c/n_0
$g(m, n; t)$	unknown function in Lanchester's equations.
G	number of groups
I	(integral) exchange ratio ($= E^{1/(1+\gamma)}$)
m	instantaneous numbers of combatant troops in Red army, usually the loser
m_0	value of m at beginning of battle
m_f	value of m at end of battle
m_c	Red casualties ($= m_0 - m_f$)
n	instantaneous numbers of combatant troops in Blue army, usually the winner
n_0, n_f, n_c	see m_0, m_f, m_c above
$N(w')$	number of battles with $w < w'$
$P(w')$	probability that Blue (or stronger) will win if $w = w'$
R_c	larger of (n_c/n_0) and (m_c/m_0)
R_f	larger of (n_0/m_0) and (m_0/n_0)
t	time or timelike quantity, or "Student" 's statistic
V_s	expected number of battles won by stronger
w	the variable used in Brown's theory to predict victory
x	log n_c/m_c, the independent variable in correlation analysis; also mz in App A
y	log m_0/n_0 or log m_0, dependent variable in correlation analysis; also n/z in App A
z	log n_0, dependent variable in correlation analysis; also $I^{1/2} = E^{1/[2(1+\gamma)]}$, parameter in Brown's equations
β	exponent describing loser's effectiveness in asymmetric Lanchester equations
γ	exponent describing both opponents' effectiveness in symmetric Lanchester equations
δ	exponent describing winner's effectiveness in asymmetric Lanchester equations
ϵ	log E in correlation analysis
χ^2	chi squared, Pearson's statistic
ρ	correlation coefficient, generally ρ_{xy}
ρ_{xy}	correlation coefficient describing dependence of x on y
$\rho_{xy,z}$	partial correlation coefficient describing dependence of x on y when z is held fixed
σ_s^2	variance of number of battles won by stronger

562 Deutsch-Französischer Krieg 1870–1871.

1870 27./11.

SCHLACHT
bei
Amiens (5.)
(Villers-Bretonneux)

(Stadt in Frankreich, Hauptstadt des Dép. Somme, 133 km nördl. von Paris).

Sieg der **Deutschen** (30.600 Inft., 4.400 Kav., 137 Gesch. = **35.000 M.**) unter Gen. d. Inf. Fh. v. Manteuffel über die **Franzosen** (23.500 Inft., 1.500 Kav., 42 Gesch. = **25.000 M.**) unter Gen. Faidherbe.

Verluste:

	230 (1 Stb. 18 Offz.)	tot	(1 Stb. 13 Offz.) 260	
	1.070 (6 „ 50 „)	verwundet	(3 „ 33 „) 1.140	
3·7% =	1.300 (7 Stb. 68 Offz.)	**Blutige Einbusse**	(4 Stb. 46 Offz.) 1.400 =	5·6%
0·8% =	300 (— 1 „)	vermißt, gefangen	(— 20 „) 2.100	8·4%
4·5% =	**1.600** (76 Offz.)	**Gesamt-Verlust**	(70 Offz.) **3.500**	**14·0%**

Verl. an Trophäen: 9 Kanonen (= 21%), 2 Fahnen.

1870 15.–27./11.

BELAGERUNG
und
EINNAHME
von
La Fère

(Stadt in Frankreich, Dép. Oise, an der Oise, 22 km nordwestl. von Laon).

Die **Deutschen** (5.000 M.) zwingen die **französische** Besatzung (2.300 M., 70 Gesch.) zur Übergabe. Die Garnison wurde kriegsgefangen.

1870 28./11.

SCHLACHT
bei
Beaune-La-Rolande (4.)

(Stadt in Frankreich, Dép. Loiret, an der Rolande, 10 km südöstl. von Pithiviers).

Deutsche — GFM. Pz. Friedrich Karl v. Preußen

Franzosen — Gen. Crouzat

Streitkräfte:

	31.000	Infanterie	56.000	
	6.000	Kavallerie	4.000	
174 Gesch.	**40.000**	**Gesamt-Stärke**	**60.000**	Gesch. 138

Verluste:

	1.000 (1 Stb. 39 Offz.)	tot und verwundet	(11 Stb. 112 Offz.) 2.200	3·7%
		vermißt, gefangen	— — 1.800 =	3·0%
2·5%	**1.000**	**Gesamt-Verlust**	**4.000** =	**6·7%**

1 Geschütz. **Verl. an Trophäen:**

Fig. 1—Sample Page from Bodart's Militär-historisches Kriegs-Lexikon, (1618–1905)[7]

INTRODUCTION

To determine the historical influence on the outcome of battle of a "force" or "constraint" as expressed in Lanchester's equations,[1] the initial and final strengths of combatants were analyzed in almost 1500 land battles, fought mostly by Europeans on European soil from 1618 to 1905.

The source of the data is Bodart's Militär-historisches Kriegs-Lexikon,[7] well known for its completeness and accuracy, from which a sample page has been reproduced as Fig. 1. The strength and casualty figures reported by Bodart were transferred to punched cards,* along with tags identifying the winners and losers, the completeness of the data, the category of the battle (see Sec 1), and the page reference in the Lexikon (where battles are listed chronologically). The designation of a given combatant as winner was Bodart's and presumably reflects the judgment of historians generally.

A set of statistical tests was chosen and, on account of the magnitude of work anticipated, programmed for an IBM 7090 computer. In Secs 2 and 3 the Lanchester equations are translated into forms suitable for use in the reduction of the data and for comparison with the data. Section 4 presents some statistical summaries requiring for their calculation no particular theory of battle, and reports of the analysis based on the deterministic and the stochastic forms of Lanchester's equations are given in Secs 5 and 6, respectively.

1. DATA

For this study all the sea battles, as well as those land battles for which the initial strength of at least one of the combatants was missing, were deleted from Bodart's list, leaving as the sample the numbers of battles reported in Table 1. To simplify the analysis those battles that could be roughly described as meeting engagements were grouped into Category I (Bodart's classes: treffen, gefecht, and schlacht) and those which were sieges, attacks on forts, etc., in Category II (Bodart's classes: belagerung, einnahme, erstürmung, kapitulation, and überfall).† Two classes of casualties were recognized: "dead and wounded" collectively and "missing and captured." If only one number was given for total losses, and the text did not indicate to the contrary, this number was recorded under the "dead and wounded" heading. Since such an aggregation of data was more likely to occur for the winner than the loser, the classification is probably in accord with reality. Although the casualty figures used in the analysis were those of the first type unless otherwise

*Copies of this card deck are available.

†These are not to be confused with Bodart's six Categories, which are defined by the total number of casualties on both sides.

indicated, many of the conclusions of this study do not lean on casualty data, and the analytical results are mainly independent of this choice.

TABLE 1

NUMBERS OF BATTLES DETAILED IN BODART

Condition	Category I, engagements	Category II, sieges	Total
Initial-strength data[a]	82	323	405
Both initial-strength and casualty data[b]	939	149	1088
Total	1021	472	1493

[a]Only initial-strength data available.

[b]Both initial-strength and casualty data available for both sides.

2. THEORY: CORRELATION ANALYSIS

Assume that each of the battles for which there are initial- and final-strength data represents a possible solution to a generalized Lanchester law:

$$\frac{dm}{dt} = -E^{-1/2}\, n^{\gamma} g(m, n; t)$$
$$\frac{dn}{dt} = -E^{1/2}\, m^{\gamma} g(m, n; t), \qquad (1)$$

where m is the instantaneous strength of the loser (Red), and n of the winner (Blue), $g(m, n; t)$ is a term symmetric under interchange of m and n, E is some power of the exchange ratio, t is a timelike variable, and γ is an exponent whose value determines the form of the law. If $\gamma = 1$ the square law obtains, and if $\gamma = 0$, the linear law.*

It can be argued that a real battle consists of many small frays. In fact, Lanchester himself so argued. Were each fray to satisfy all the requirements of the linear law or square law, while collectively they were to display a spectrum of values of the parameters in Lanchester's equations, one would expect the overall effect to be "intermediate" to a pure case of $\gamma = 0$ or $\gamma = 1$. If this were in fact true, then an appropriate method of averaging such a set of data as Bodart's should yield an "effective" value of γ somewhere between 0 and 1. This section and the one following are devoted to methods of deriving an effective value of γ, whatever the circumstances that generate it. (The question of the existence of an effective value of E is discussed in Sec 5.†)

Since (a) the durations of the conflicts were not well reported and (b) the relation between the variable t and actual time involves additional unknown em-

*Morse and Kimball[10] assume $g = 1$ for both laws. Weiss[2] takes $g = 1$ for the square law, $g = mn$ for the linear law.

†It can also be argued that this heterogeneity of battle implies so complicated a generalization of Lanchester's equations as to cast serious doubt on their utility. See Snow[3] and Smith.[11]

pirical parameters, time has not been used as a datum.* Rather this integral of Eq 1 was used:

$$E = \frac{n_0^{1+\gamma} - n_f^{1+\gamma}}{m_0^{1+\gamma} - m_f^{1+\gamma}}, \tag{2}$$

which is independent both of t and of the function β. (The subscript 0 refers to initial strengths, subscript f to final strengths.)

One method of averaging involves an approximation to Eq 2. It may be rewritten

$$n_0^{1+\gamma}\left[1-(1-f_n)^{1+\gamma}\right] = E m_0^{1+\gamma}\left[1-(1-f_m)^{1+\gamma}\right] \tag{3}$$

where $f_n = 1 - \dfrac{n_f}{n_0} = \dfrac{n_c}{n_0}$

$f_m = 1 - \dfrac{m_f}{m_0} = \dfrac{m_c}{m_0}$

n_c, m_c = casualties to Blue and Red, respectively

Since in 90 percent of the battles of Category I both f_m and f_n were less than $\frac{1}{3}$ (see Table 6) the left- and right-hand sides of Eq 3 can be expanded by the binomial theorem, keeping only the leading terms:

$$n_c n_0^{\gamma}(1 + \ldots) = E m_c m_0^{\gamma}(1 + \ldots)$$

or

$$\log \frac{n_c}{m_c} = \log E + \gamma \log \frac{m_0}{n_0} + \ldots$$

or with an obvious substitution:

$$x = \epsilon + \gamma y. \tag{4}$$

Equation 4 was used to derive γ by linear regression. The results are reported in Sec 5.

To express the possibility that winner and loser may fight differently in a Lanchester-type situation, Eq 1 may be modified as follows:

$$\begin{aligned} \frac{dm}{dt} &= -E^{-\frac{1}{2}} n^{\delta} \\ \frac{dn}{dt} &= -E^{\frac{1}{2}} m^{\beta}. \end{aligned} \tag{5}$$

*Assuming for the moment that $E = 1$, let $m = r \cos \theta$ and $n = r \sin \theta$. Then Weiss' formulation of the two laws leads to

$$m\frac{dn}{dt} - n\frac{dm}{dt} = r^2 \frac{d\theta}{dt} = k(mn)^{1-\gamma}$$

indicating that the rate at which area is swept out by the radius drawn from the origin to the point with Cartesian coordinates (m, n) is independent of the time in the square-law case, but is a decreasing function of the time in the linear case. Knowledge of the casualties at a time that splits the battle into known fractions is thus the basis of a sharper test for γ.

Then the analog of Eq 2 is

$$E = \frac{1+\beta}{1+\delta} \frac{n_0^{1+\delta} - n_f^{1+\delta}}{m_0^{1+\beta} - m_f^{1+\beta}}$$

and of Eq 4 is

$$\log \frac{n_c}{m_c} = \log E + \beta \log m_0 - \delta \log n_0$$

or by an obvious substitution:

$$x = \epsilon + \beta y - \delta z \tag{6}$$

The results of the regression analysis using Eq 6 are also reported in Sec 5.

3. THEORY: STOCHASTIC LANCHESTER EQUATIONS

A more elegant treatment results from dropping the assumption of Sec 2 and adopting the weaker hypothesis that these battles obey a stochastic form of Lanchester's equations. Such equations are derived and discussed by Morse and Kimball[10] and from a somewhat more fundamental point of view by Brown.[12] It is clear that Eq 1 (or Eq 5 for that matter) is independent of the choice of scale for m and n. These symbols can be interpreted variously as numbers of men, companies, divisions, etc., without altering the form or superficial meaning of the answers; the equations are "essentially" homogeneous in m and n.* In the equations as treated by Morse and Kimball, m and n are stochastic variables whose dispersions are related to their magnitudes, and the solutions to the equations are the expected numbers of survivors and the dispersions of these numbers. If m_0 and n_0 are less than 10, say, these dispersions seem to be in conformity with experience; on the other hand if m_0 and n_0 are as large as 10^4 to 10^5 the dispersions are vanishingly small. (The argument in App A shows that the solutions of the set of equations that are the continuous-variable analog of the stochastic equations yield zero dispersion about the solutions to Eq 1.)

It is clear then that in any application of the stochastic equations a scaling rule must be used. An appealing choice of a scale factor is

$$SF = \sqrt{m_0 + n_0} \tag{7}$$

times some quantity that is nearly unity. In essence this is what Brown does in deriving approximate solutions to the stochastic equations, valid in the limit of large m_0 and n_0. His conclusion is that for fixed m_0 and n_0 (m standing for Red, n for Blue) the probability that Blue wins is given by

$$\left.\begin{aligned} &P(w) = 1 \text{ if } w > w_2 \\ &P(w) = \frac{\int_{-w_1}^{w} \exp(-\tfrac{1}{2}t^2)\,dt}{\int_{-w_1}^{w_2} \exp(-\tfrac{1}{2}t^2)\,dt} \text{ if } -w_1 \le w \le w_2 \\ &P(w) = 0 \text{ if } w < -w_1, \end{aligned}\right\} \tag{8}$$

*Admittedly the definition of a casualty is less obvious when m and n count companies or divisions. Are 10 companies that have taken 30 percent casualties as effective as 7 unscathed?

where

$$\left.\begin{aligned} w &= (n_0/z - m_0 z)\left[(1+2\gamma)/(m_0+n_0)\right]^{1/2} \\ w_1 &= z\left[(1+2\gamma)(m_0+n_0)\right]^{1/2} \\ w_2 &= -\frac{1}{z}\left[(1+2\gamma)(m_0+n_0)\right]^{1/2} \\ z &= (E)^{\frac{1/2}{1+\gamma}}. \end{aligned}\right\} \qquad (9)^*$$

If w_1 and w_2 are large enough we have approximately

$$P(w) = (2\pi)^{-1/2} \int_{-\infty}^{w} \exp(-\tfrac{1}{2}t^2)\, dt. \qquad (10)$$

It is worth while to observe that Eq 2 can be rewritten

$$(n_0/z)^{1+\gamma} - (m_0 z)^{1+\gamma} = (n/z)^{1+\gamma} - (mz)^{1+\gamma},$$

so that w predicts, with probability $P(w)$ or $1-P(w)$, the outcome that Eq 2 predicts with certainty. $P(w)$ can be computed if m_0, n_0, γ, and E are known. If the battle has been a true Lanchester battle then E is given, at least approximately, by Eq 2. $P(w)$ can then be computed if m_0, m_f, n_0, n_f, and γ are known. It is assumed for the purposes of the following analysis that one is justified in computing E from Eq 2. Then γ is the only free parameter to be determined. (To appreciate a later result, note that the expressions for w, w_1, and w_2 all require that γ be greater than $-.5$.)

(Another approach, not adopted here, is to take m_0 and n_0 in Eq 9 to stand for the strengths after known numbers of casualties have been suffered, but to compute z from the initial strengths and these casualties as before. P would then have been given the meaning of the probability of a Blue victory if the combatants had not decided to quit at this particular time. There are semantic difficulties in pursuing this course.)

Assume for the moment that Blue (= the ultimate winner) is always the stronger and introduce the function

$$\begin{aligned} N(w') &= \text{number of battles with } w < w' \\ N(-\infty) &= 0 \\ N(+\infty) &= N, \text{ the total number of battles in sample.} \end{aligned}$$

Then V_s, the expected number of victories by the stronger, is given by

$$V_s = \int_{-\infty}^{\infty} P(w) \frac{dN(w)}{dw}\, dw \qquad (11)$$

since this is an example of a Bernoulli process if one neglects in particular the serial correlation of victories in battles of a given war. The variance of the predicted number of victories is simply

$$\sigma_s^2 = \int_{-\infty}^{\infty} P(w)[1-P(w)] \frac{dN(w)}{dw}\, dw. \qquad (12)$$

*Equation 9 actually constitutes an extension of Brown's argument to the case of an arbitrary value of γ. Brown assumed that $\gamma = 1$.

For those battles in which Red (m the loser) is stronger $P(w)$ is replaced by $P(-w)$. This leads to ambiguity where $m = n$ but $w \neq 0$ and $P \neq \frac{1}{2}$. In this latter case one may either take $P = \frac{1}{2}$ or disregard such battles (about 3 percent).

For an actual set of battles V_s and σ_s^2 will be computed from

$$V_s = \sum_{i=1}^{N} Q_i, \tag{13}$$

and

$$\sigma_s^2 = \sum_{i=1}^{N} Q_i(1-Q_i), \tag{14}$$

where

$$Q_i = \begin{Bmatrix} P(w_i) \\ \frac{1}{2} \\ P(-w_i) \end{Bmatrix} \text{ if } \begin{Bmatrix} n > m \\ n = m \\ n < m \end{Bmatrix}.$$

The sums converge to the corresponding integrals of Eqs 11 and 12 in the limit of large numbers. For these sums there are direct tests.

If the N battles are broken into G groups, each of size N/G, one can compute V_{sg} and σ_{sg}^2 ($g = 1, 2, \ldots, G$). A_{sg}, the actual number of battles won by the stronger combatant in the battles of the gth group, can be taken from Bodart. These quantities can be combined as the parameter

$$\chi^2 = \sum_{g=1}^{G} \left(\frac{A_s - V_s}{\sigma_s} \right)_g^2 \tag{15}$$

which should obey a chi-square distribution in G degrees of freedom.* One can also observe how V_s, σ_s^2, and χ^2 vary with γ. The significant test uses data so organized that the battles of a given group have (nearly) the same force ratio R_f (see Sec 2).

The argument for the stochastic equations is that they allow for fluctuations about the "solution path" of the deterministic equations. The validity of this approach depends in turn on the validity of an additional unstated hypothesis, viz., that military action during any of these excursions tends to return the variables to values on the solution path. This restoration would be the effect of the force alluded to in the title and the "Introduction." If an army is able to read in Lanchester's equations its imminent defeat, it will surely do everything in its power to modify the parameters and the outcome of battle. To the extent that the stochastic equations remain unsupported in this analysis, the writer feels the hypothesis of stability is also questionable.

4. FORCE RATIO, INCIDENCE OF BATTLE, AND VICTORY

In a sense Tables 2 and 3 contain the results of this study. For each Category (meeting engagements or sieges) the battles in Table 2 have been grouped

*See Cramér,[13] Chap. 30. The value $G = 1$ is not excluded. Readers may expect to find V_s rather than σ_s^2 in the denominator. The conventional derivation of χ^2 assumes that A_s and V_s describe a Poisson process, in which case the variance of V_s is indeed V_s. It is precisely for the reason that this a Bernoulli process that this different variance appears.

according to force ratio R_f.*

$$R_f = \max (n_0/m_0,\ m_0/n_0).$$

The difference in these distributions is sufficient to merit the use of the two separate categories. The fact that the numbers for Category I are well approximated by some form of exponential curve was first, but independently, noted by Smith and Donovan.[8]

TABLE 2

HISTORICAL INCIDENCE OF BATTLE VS FORCE RATIO

Range of force ratio R_f	Category I[a]	Category II[b]
1–1.5	473	64
1.5–2	251	51
2–2.5	122	51
2.5–3	58	38
3–4	56	69
4–5	30	37
5–6	9	38
6–7	13	22
>7	9	102

[a]1021 battles.
[b]472 battles.

TABLE 3

FREQUENCY OF VICTORY BY NUMERICALLY INFERIOR COMBATANT

Range of force ratio R_f	Category I		Category II	
	Frequency	Battles fought[a]	Frequency	Battles fought[a]
1–1.5	.42	(473)	.39	(64)
1.5–2	.35	(251)	.28	(51)
2–2.5	.25	(122)	.18	(51)
2.5–3	.43	(58)	.13	(38)
3–4	.27	(56)	.25	(69)
4–5	.37	(30)	.19	(37)
5–6	.11	(9)	.05	(38)
6–7	.23	(13)	.05	(22)
>7	.33	(9)	.09	(102)

[a]Values in parentheses are from Table 2.

The contents of Table 2 are repeated in Table 3 together with the frequency of victory by the numerically inferior† combatant. Taking into account the reliability of the data source the only possible conclusion is that for Category I battles and for force ratios no more disparate than 5 to 1 the outcome of battle is independent of the force ratio. Though Blue be numerically weaker† the odds

*Note that this definition is perhaps unique to this paper for the reason that the attacker was not identified.

†Throughout this paper the words "weaker" and "inferior" are always to be understood in terms of numbers of troops.

in favor of a Blue victory (3 to 5) are independent of whether Red's numerical superiority is 1.1 to 1 or 5 to 1.

To test the universality of these results the categories were subdivided according to the combined strengths of opposing forces and the year of the battle. Table 4 defines 12 groups and shows the number of battles in each;

TABLE 4

MAJOR GROUPINGS OF BATTLES

Force or year	Category	
	I, engagements	II, sieges
By total forces ($m_0 + n_0$)		
2,500–25,000	383	252
25,000–50,000	346	134
50,000–640,000	292	86
Total	1021	472
By year of occurrence		
1618–1792	338	241
1792–1815	469	178
1815–1905	214	53
Total	1021	472

Table 5 contains battle-incidence and victory data for each. The only deviations from the patterns in Table 3 occur for Category II: for the medium-sized forces (Table 5e) and for the post-Napoleonic period (Table 5c). In both instances the frequencies of victory to the weaker are roughly constant (like the general results for Category I) although the value of this constant is less (odds of about 1 to 3).

Another test used the casualty ratios f_m and f_n defined under Eq 3. The larger of these two numbers R_c was recorded for each battle, and Table 6 shows the number of battles in Category I for which R_c is less than a specified fraction R. For those battles for which R_c was in excess of 1/3, Table 7 shows the incidence of battle and frequency of victory by the weaker combatant, grouped as before according to force ratio R_f.

If finally in the light of the results in Table 5c (Cat. II), only the 84 battles in Category I that occurred later than 1870 (beginning with the Franco-Prussian war) are examined, it is found that even that bias against the weaker combatant reported earlier has disappeared. For force ratios less disparate than 4 to 1 the odds were 50–50 (see Table 8). Although it will be recognized that this result could be accounted for by the swing of just seven battles, the implications for modern-day battle are sufficiently interesting to warrant similar investigations of combat since 1905.

Inasmuch as force ratio seems to play a minor role in the decision of battle it should not be expected that Lanchester's equations, which are built around force ratio, will provide a nice determinant of winner and loser. Contrariwise, working the equations backward from the outcome of battle to the parameters of the theory is likely to yield unsatisfactory results. The next two sections bear this out.

TABLE 5

ALL BATTLES WITH AND WITHOUT CASUALTY DATA

Range of force ratio	a: 1618–1792		b: 1792–1815		c: 1815–1905		d: <25,000 men		e: 25,000–50,000 men		f: > 50,000 men	
	Number	Probability	Number	Probability	Number	Probability	Number	Probability	Number	Probability	Number	Probability
					Category I							
1–1.5	158	.44	226	.38	97	.49	155	.42	162	.43	161	.42
1.5–2	84	.46	115	.25	53	.32	94	.26	81	.33	77	.44
2–2.5	40	.48	53	.19	24	.54	52	.40	45	.27	23	.39
2.5–3	17	.53	27	.26	12	.50	24	.29	20	.35	9	.78
3–4	18	.33	27	.33	12	.33	32	.34	16	.25	12	.42
4–5	8	.63	13	.31	6	.17	11	.46	12	.42	5	.40
5–6	6	.33	2	.50	3	0	5	0	2	0	1	0
6–7	6	.33	2	.50	3	0	5	.60	5	.20	2	0
> 7	1	0	4	.25	4	.50	4	.50	3	.33	2	0
					Category II							
1–1.5	22	.41	36	.50	9	.22	39	.39	17	.41	10	.50
1.5–2	18	.33	28	.21	10	.20	32	.28	15	.27	11	.18
2–2.5	13	.06	22	.27	9	.22	32	.16	9	.22	8	.25
2.5–3	17	.06	12	0	4	.25	21	.05	4	.25	10	.10
3–4	42	.29	23	.17	8	.25	45	.18	17	.41	8	.25
4–5	20	.25	13	0	4	.25	25	.16	12	.25	1	0
5–6	20	.15	10	.10	5	0	15	.07	16	.13	4	0
6–7	16	.13	4	0	1	0	12	.08	3	.33	5	0
> 7	68	.07	30	.10	3	.33	31	.07	41	.12	29	.07

TABLE 6

HISTORICAL INCIDENCE OF CASUALTY RATIOS

(Category I)

Critical casualty ratio R	Fraction of battles with R_c[a] less than R
.05	.13
.075	.25
.10	.41
.125	.50
.20	.73
.25	.82
.33	.90
.50	.98

[a] $R_c = \max\left(\frac{n_c}{n_0}, \frac{m_c}{m_0}\right)$

TABLE 7

INCIDENCE OF BATTLE AND FREQUENCY OF VICTORY BY WEAKER COMBATANT

Range of force ratio	Battles with $R_c \geq 1/3$	Frequency of victory by weaker combatant
1–1.5	25	.24
1.5–2	18	.28
2–2.5	13	.08
2.5–3	5	.20
3–4	4	0
4–5	5	0
5–6	0	0
6–7	1	0
> 7	4	.50

TABLE 8

84 BATTLES, 1870–1905

(Category I)

Range of force ratio	Battles	Frequency of victory by weaker side	Frequencies averaged over extended ranges
1–1.5	32	.59	.52
1.5–2	18	.39	.50 .52
2–2.5	12	.42	.49
2.5–3	6	.67	
3–4	6	.17	.44
4–5	2	0	
5–6	1	0	
6–7	4	0	
> 7	3	.33	

5. RESULTS OF CORRELATION ANALYSIS

We now test the hypothesis that a set of N battles may be effectively "described" by Eq 4 with parameters ϵ and γ or by Eq 6 with parameters ϵ, β, and δ. Assuming that the deviations in x, y, and z are normally distributed, x is treated as the independent variable and the N values of x, y, and z are subjected to a least-squares analysis that yields two equations in ϵ and γ or three in ϵ, β, and δ. At the same time the correlation coefficients ρ_{xy}, ρ_{xz}, and ρ_{yz} are computed. In case γ is sought (Eq 4) the value of $(\rho_{xy})^2$ is the fraction of the variance of x that is explained by the hypothesized equation and the variations of y. When solving for values of β and δ it is better to look at the partial correlation coefficients $\rho_{xy,z}$ and $\rho_{xz,y}$, where for example,

$$\rho_{xy,z} = \frac{\rho_{xy} - \rho_{xz}\rho_{yz}}{[(1-\rho_{xz}^2)(1-\rho_{yz}^2)]^{1/2}}$$

describes the part of the variation in x that is due purely to y, assuming z is held constant (see Hald[14] for complete details).

Analysis using Eq 4 shows the extent to which the casualty ratio is related to the force ratio, and that using Eq 6 shows the extent to which the casualty ratio is related separately to the loser's and the winner's strengths. Should $\rho_{xy,z}$ and $\rho_{xz,y}$ be individually considerably different from zero but add up to zero, one then knows that loser and winner are contributing equally to the casualty ratio. Should they be individually near zero then one knows, and the values of β and δ will confirm, that the casualty ratio is sensibly independent of initial strengths. The linear law would predict this result.

TABLE 9

RESULTS OF CORRELATION ANALYSIS USING EQ 4[a]

Subcategory	Category I			Category II		
	Battles[b]	γ	ρ	Battles[b]	γ	ρ
By total forces						
2,500–25,000	330	–.51	.35	51	–.62	.60[c]
25,000–50,000	323	–.35	.27	49	–.45	.59[c]
50,000–640,000	285	–.42	.30	49	–.43	.57[c]
By time frame						
1618–1792	292	–.50	.32	75	–.61	.65[c]
1792–1815	437	–.35	.24	48	–.27	.48
1815–1905	208	–.41	.35	26	–.61	.52[c]
By force ratio						
1–1.5	476	–.38	.12	27	–.87	.27
1.5–3	371	–.4	.29	54	–.37	.32
> 3	91	–.55	.58[c]	70	–.46	.60[c]

[a] $\log \frac{n_c}{m_c} = \log E + \gamma \log \frac{m_0}{n_0}$

$x = \epsilon + \gamma y$

[b] Any slight disparity between the numbers in this column and comparable data elsewhere is real and results from detection of errors in basic data as analysis proceeded.

[c] Correlation coefficient over .5.

Table 9 summarizes the results for both Categories, using Eq 4. For all subcategories examined, and for the battles as a whole, γ was found to be not 0 or 1 or somewhere between, but lying between −.27 and −.87 with typical values around −.5, especially when the correlation coefficient was highest. Inasmuch as the limits of integration in Eq 9 all contain $(1 + 2\gamma)^{1/2}$ as a factor, it is clear that for mathematical reasons one hopes that $\gamma > -\frac{1}{2}$. Military theorists should be discouraged to find $\gamma < 0$, for in this range the results seem to imply that if the Lanchester formulation is valid the casualty-producing power of troops increases as they suffer casualties.

Helmbold[1] has discovered for a small set of battles the same effect. Since he assumes the validity of the square law he interprets this result, not as a denial of the square law, but rather as a demonstration of the dependence of E on the force ratio. Inasmuch as he has not proved that the square law describes his battles his explanation is not compelling.

The results of the correlation analysis using Eq 5 are summarized in Table 10 for Category I battles and in Table 11 for those of Category II. The

TABLE 10

RESULTS OF CORRELATION ANALYSIS OF CATEGORY I BATTLES, USING EQ 5[a]

Subcategory	Battles	β	δ	ρ_{xy}	ρ_{xz}	ρ_{yz}[b]	$\rho_{xz,y}$
By total forces							
2,500–25,000	330	−.37	−.62	.33	[illegible]	[illegible]	.39
25,000–50,000	323	−.23	−.52	[illegible]	[illegible]	−.16	.32
50,000–640,000	286	−.37	−.49	.20	[illegible]	[illegible]	[illegible]
By time frame							
1618–1792	292	−.42	−.59	.24	.01	−.26	[illegible]
1792–1815	437	−.18	−.55	.40	.22	−.13	[illegible]
1815–1905	210	−.39	−.52	.25	−.03	−.32	[illegible]
1870–1905	84	−.50	−.60	.21	−.15	−.36	.48
By force ratio							
1–1.5	477	−.27	−.54	.30	.26	−.09	.17
1.5–3	371	−.33	−.54	.27	.05	−.23	.35
> 3	91	−.48	−.59	.48	−.29	−.38	.53[c]
By exchange ratio[d]							
$I' = 1.225$	76	.20	.20	−.08	.04	.42	−.43
$I' = 1.5$	260	.13	.09	.12	.29	.20	−.13
$I' = 1.838$	410	.01	−.05	.16	.14	.02	.07
By casualty ratio[d]							
0–.127	469	−.54	−.62	.22	−.50	−.39	.43
.128–.333	395	−.27	−.51	.32	.11	−.19	.35
> .334	75	−.29	−.56	.34	.12	−.17	.36
All battles	939	−.31	−.57	.33	.11	−.22	.37
n_c changed randomly	939	−.31	−.57	.33	.11	−.21	.37
n_c up 10 percent, m_c down 10 percent	939	−.31	−.57	.33	.11	−.22	.37
Use all casualties	939	−.21	−.52	.32	.16	−.13	.31

[a] $\log \dfrac{n_c}{m_c} = c + \beta \log m_0 - \delta \log n_0$

$x = c + \beta y - \delta z$

[b] The minus sign is proper in this column if $\beta < 0$.

[c] Correlation coefficient over .5.

[d] See Sec 5 for explanation of exact criterion.

TABLE 11

RESULTS OF CORRELATION ANALYSIS OF CATEGORY II BATTLES, USING EQ 6

Subcategory	Battles	β	δ	ρ_{xy}	ρ_{xz}	$\rho_{xy,z}$	$\rho_{xz,y}$
By total forces							
2,500–25,000	51	−.39	−.87	.58[a]	−.28	−.33	.60[a]
25,000–50,000	49	−.40	−.53	.45	−.45	−.43	.43
50,000–640,000	49	−.40	−.57	.33	−.48	−.50[a]	.37
By time frame							
1618–1792	75	−.59	−.62	.43	−.39	−.55[a]	.57[a]
1792–1815	48	−.21	−.34	.40	−.30	−.32	.41
1815–1905	26	−.50	−.70	.42	.06	−.43	.57[a]
By force ratio							
1–1.5	25	−.69	−.94	.37	.31	−.23	.31
1.5–3	54	−.38	−.37	.06	−.16	−.32	.28
> 3	70	−.35	−.57	.52[a]	−.32	−.38	.55[a]
By exchange ratio							
$I' = 1.225$	6	−.04	−.05	.17	.16	−.07	.09
$I' = 1.5$	20	.21	.16	.15	.31	.41	−.32
$I' = 1.838$	34	.10	.15	−.18	−.10	.15	−.21
By casualty ratio							
0–.127	33	−.46	−.46	.12	−.46	−.67[a]	.55[a]
.128–.333	63	−.39	−.57	.46	−.23	−.40	.54[a]
≥ .334	53	−.45	−.67	.51[a]	−.31	−.47	.60[a]
All battles	142	−.44	−.55	.39	−.30	−.49	−.54[a]

[a]Denotes correlation coefficient over .5.

arrangement of data is the same as in Table 9, and direct comparisons can be made. It will be seen that the values of β and δ bracket the values of γ in Table 9, but that $\delta \approx \beta - .2$ and is almost always less than −.5. On the basis of the values of the ρ's, one would say that Eq 6 does not give a significantly better fit to the data than Eq 4; on the other hand the consistency of sign of $\delta - \beta$ is a strong indication of a real effect.

The next set of tests required a selection of those battles for which the exchange ratio was nearly unity. With γ unknown it was not desirable to bias the sample in favor of an erroneous choice of γ. Two precautions were taken:

(a) Since Lanchester is being tested, in terms not of Eq 1 but of the time-independent Eq 2, any exchange ratio used in this analysis should be that pertinent to Eq 2. Combatants will maintain the same force ratio throughout the battle if $E = (n_0/m_0)^{1+\gamma}$ or $(n_0/m_0) = E^{1/(1+\gamma)}$. Accordingly it is appropriate to define

$$I = E^{1/(1+\gamma)}$$

and name it the integral exchange ratio, coming as it does from an integral rather than from a differential equation. I has an advantage over E in that it is generally less sensitive than E to changes in γ.

(b) For each battle two values of E were computed—E_0 and E_1—using Eq 2 and a value of γ equal to the subscript. From them I_0 and I_1 were computed in turn. A battle was rejected unless both I_0 and I_1 fell in the range $(I', 1/I')$, where I' was an arbitrarily chosen constant.

The test was performed with three different values of I':

$$(1.5)^{1/2} \approx 1.225;\ 1.5;\ (1.5)^{3/2} \approx 1.838,$$

and the columns so labeled in the tables refer to the results of these tests.

Since only for Category I was the sample of any size, only those battles are discussed. It will be observed that:

(a) β and δ are now the same within experimental error.

(b) The average of β and δ increases as I' becomes smaller and that even for the largest I' used this average is essentially zero. These groups constitute the only ones for which the writer has succeeded in extracting a positive value of the exponent.

(c) The partial correlation coefficients increase systematically as I' is decreased, showing that a real effect is being isolated. According to the t-test the correlation is significant at the 0.1 percent level for the smallest value of I' (i.e., when data are drawn from a population that has a Gaussian distribution of errors, in 999 samples out of 1000 a true correlation coefficient of zero would result in an observed value less than the value reported here).

By selecting battles according to values of the integral exchange ratio, the computer has thus isolated a subset of Category I (almost 50 percent of it) for which the effective value of γ is about zero. These are battles in which the two opponents are nearly evenly matched (values of I' near unity). For these battles more information is available. Table 12 shows in the format used earlier the incidence of battle and the frequency of victory by the weaker side. Note that in contrast to Category I as a whole the weaker side wins much more rarely as

TABLE 12

INCIDENCE OF BATTLE AND PROBABILITY OF VICTORY BY WEAKER SIDE

Range of force ratio	I'					
	1.225		1.5		1.838	
	Battles	Probability	Battles	Probability	Battles	Probability
			Category I			
1–1.5	72	.39	166	.40	228	.39
1.5–2	4	0	74	.16	113	.25
2–2.5	0	—	14	0	40	.18
2.5–3	0	—	3	0	12	.17
3–4	0	—	3	0	15	0
4–5	0	—	0	—	2	.50
			Category II			
1–1.5	6	.17	10	.20	11	.18
1.5–2	0	—	6	.17	12	.25
2–2.5	0	—	2	0	6	0
2.5–3	0	—	2	0	4	0
3–4	0	—	0	—	1	0
4–5	0	—	0	—	0	

the adverse force ratio increases. For the larger values of I' the incidence of battle follows the general behavior of Category I as a whole, but for $I' \doteq 1.225$ all force ratios are less than 2 and the overwhelming number less than 1.5. They are not typical of all battles with small force ratios. The identities of these 76 battles were obtained and additional informat on about them extracted from Bodart. In less than 10 percent did the bloody casualties suffered by either side exceed 20 percent. More importantly, almost none of these battles would be called a "significant" one from a historical point of view (see App B).

For the 939 − 410 = 529 battles for which I (I_0, or I_1, or both) is outside the range defined by $I' = 1.838$ the correlation between x and y is so small that it was assumed that $x = \epsilon - \delta z$. In this case it was found that $\delta = -.328$ with a correlation coefficient $\rho_{xz} = .29$. By the t-test this result is significant at a level beyond 0.01 percent. How is it possible that a set of battles for which $\beta \approx \delta \approx 0$ can combine with some for which $\delta \approx -1/3$ to give a composite with still more negative values of β and δ? The answer may lie in the lack of homogeneity in appropriate values of ϵ (or $E = e^{\epsilon}$). For the 76 battles $E = .967$, for the 410 battles $E \approx 3/5$, and for the 529 excluded battles $E \approx 1/65$.

Further to this point, for each battle the value of γ was computed (Eq 2), assuming $E = 1$. This was done by iteration, starting from a trial value of $\gamma = 0$ and terminating when the unsigned displacements of two successive iterates were less than .01 and not increasing (or in a few cases if the process failed to converge). Table 13 summarizes these calculations for battles of Category I.

TABLE 13

DISTRIBUTION OF γ FOR $E = 1$

(Category I)

$\lvert\gamma\rvert$	For $\gamma < 0$	For $\gamma > 0$	$\lvert\gamma\rvert$	For $\gamma < 0$	For $\gamma > 0$
0–.1	←44→		1–2	113	96
.1–.2	14	15	2–3	53	61
.2–.3	15	15	3–4	29	31
.3–.4	14	17	4–5	18	15
.4–.5	14	21	5–6	16	13
.5–.6	20	17	6–7	7	14
.6–.7	18	18	7–8	8	4
.7–.8	13	15	8–9	2	2
.8–.9	18	14	9–10	5	2
.9–1.0	12	12	> 10	0	29

Although the computer print-out contains the information necessary to get least-squares values for log E, only a few values of E thus derived have been computed and reported because the process of solving for logarithms by least squares generally leads to very poor values of the antilogarithms.

Data on two other types of tests are contained in Tables 10 and 11. Let the upper casualty ratio

$$R_c = \max\left(\frac{n_c}{n_0}, \frac{m_c}{m_0}\right)$$

be computed for each battle (see Table 6). The values of β and δ for three ranges of R_c are reported.

Finally, modifications in Bodart's casualty numbers were made to test the sensitivity of the results to errors in the reports. In one test n_c was multiplied by the factor

$$1 + 0.2n_r,$$

where n_r was a random number generated for each battle by a rule (the numbers tabulated by Suhler[15] were used) that assured that the set of n's was normally distributed with zero mean and unit variance, truncated at $|n| = 4$. In another test all n_c's were decreased by 10 percent and all m_c's increased 10 percent. In the final test and in this one instance all casualties—dead, wounded, missing and captured—were included in the analysis.

6. RESULTS OF STOCHASTIC ANALYSIS

Figure 2 shows a histogram of the observed distribution of values of w with the results for $\gamma = 0$ superposed on those for $\gamma = 1$. For the construction of Fig. 2, n always stands for the winner and m the loser in Eq 9. As a result the successes of prediction are registered to the right of the abscissa $w = 0$, the errors of prediction to the left. A few values beyond -100 and $+200$ have not been illustrated. For the Category I battles summarized in Fig. 2, Brown is 76.4 percent correct in predicting the outcome of battle if $\gamma = 1$, 77.8 percent correct for $\gamma = 0$. There is little to choose between $\gamma = 0$ and $\gamma = 1$ on this basis. The center of gravity $\bar{w}(\gamma)$ has values of 55 and 74 for $\gamma = 1$ and 0, respectively, and achieves its maximum value of 85 at $\gamma = -.32$.

Since $P(w)$ is essentially zero for $w < -5$ and essentially unity for $w > +5$ it follows that to the scale of Fig. 2 a plot of $P(w)$ becomes essentially the unit step function with the jump at $w = 0$ (see Fig. 3). [This is another way of saying that in spite of the scale factor introduced by Brown (see Sec 3 of this paper) the predictions of outcome by stochastic and deterministic forms of Lanchester's equations are almost identical.]

To this approximation V_s and σ_s^2 become

$$\begin{aligned} V_s &= \int_{-\infty}^{0} 0 \cdot dN(w) + \int_{0}^{\infty} 1 \cdot dN(w) \\ &= N(\infty) - N(0) \\ &= \text{number of battles with } w > 0, \end{aligned}$$

and

$$\begin{aligned} \sigma_s^2 &= \tfrac{1}{4}\,[N(0+) - N(0-)] \\ &= \tfrac{1}{4}\,(\text{number of battles with } w \text{ exactly zero}). \end{aligned}$$

Thus Brown, however right or wrong, is rarely in doubt about the outcome of battle.

The 20 to 25 percent erroneous predictions are largely compensatory, as a more detailed analysis shows. For the usual 939 battles of Category I, it is found that, using Eqs 13 and 14, Brown "predicts" 621 victories by the stronger

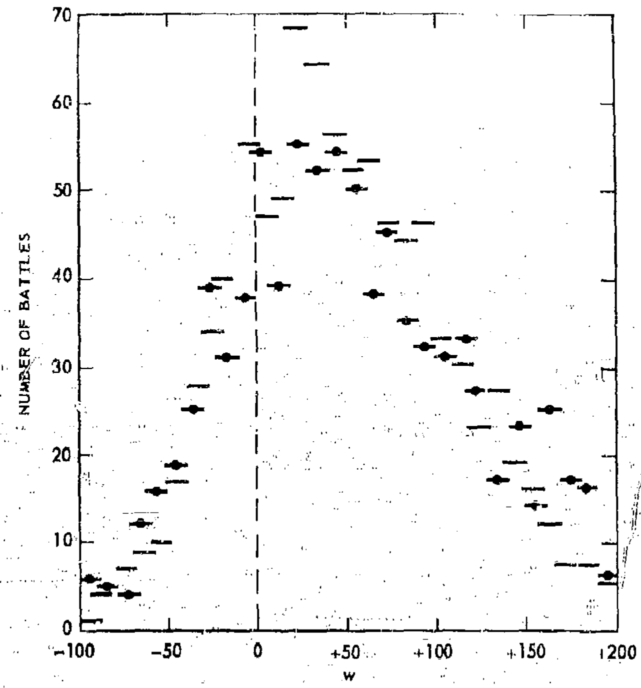

Fig. 2—Histogram of the Observed Distribution of Values of w

$\bullet$ $y = 0$ — $y = 1$

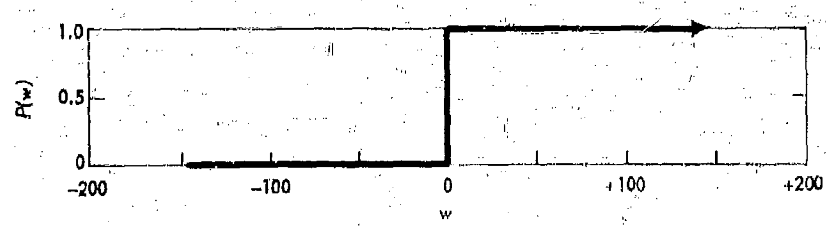

Fig. 3—Plot of $P(w)$

with a standard deviation (σ_s) of 3.2 battles; 596 actually were so won. Similarly, for 149 battles of Category II, 123 victories were predicted with a standard deviation of 1.3 battles; 121 occurred.

Table 14 shows a more detailed picture of prediction and actuality, giving the numbers for each of 31 groups of Category I and 12 groups of Category II. The battles were arranged in order of value of R_f, and the groups were formed in sequence from this list. Table 14 is thus a variant of Table 3 with different ranges of R_f. The correspondence between prediction and outcome is remarkable, especially for Category II. The coefficient of correlation between V_s and A_s is $\rho_{VA} = 0.885$, the square of which is 0.79 and the same as the fraction of correct predictions. In the case of Category I, ρ_{VA} is only 0.436, because there is a real variation of prediction with force ratio.

Brown's choice of scale factor is not unique. One can be constructed that is closer in form to what would be expected on esthetic grounds from Lanchester's equations: $SF = (n_0/z) + m_0 z)^{1/2}$.* Then w is

$$\left(\frac{n_0}{z} - m_0 z\right) \left[\left(1 + 2\gamma\right) \Big/ \left(\frac{n_0}{z} + m_0 z\right) \right]^{1/2},$$

and w_1 and w_2 merge to the common value

$$\left[\left(1 + 2\gamma\right)\left(\frac{n_0}{z} + m_0 z\right)\right]^{1/2}$$

There is a slight improvement in the range of force ratio over which prediction and observation are in accord for Category I.

For the values of V_s, A_s, and σ_s given above and $G = 1$ (one degree of freedom) one finds a χ^2 of 62 for Category I and of 3.3 to 3.8 for Category II. The probability associated with χ^2 in this latter range lies between 5 and 10 percent, and the dispersion of Category II results is not significant, though the dispersion of Category I data requires another explanation. (See Sec 3, first paragraph.)

If σ_s^2 is computed for each of the groups defined in Table 14, χ^2 becomes much larger on account of the distribution of the w's; for several of the groups the value of $P(1 - P)$ for every battle is less than 10^{-4}. When the groups are aggregated χ^2 may be reduced markedly, but not below the values quoted above.

It can be said that the step-function behavior of $P(w)$ explains much of the results in Table 3 but not the outstanding feature of Category I battles. The results for Category II encountered in Tables 3, 9, and 14 make more sense than those for Category I. The explanation may be simply that there is rarely any doubt about the outcome of an attack on a fort; history and Brown are very likely to agree in these cases. Perhaps the outcome of some of the Category I battles were better called indecisive, a classification here ignored. If these draws were eliminated, the two Categories might appear much more alike. Possibly the real reason for the different behavior of the two Categories is that, without any special care taken to ensure this result, Blue is almost always the attacker in Category II. It is highly unlikely that the attacker in these situations was the occupant of the fortified position, sallying forth against a

*One advantage of this choice is that it is then possible to construct an analogous stochastic theory around the asymmetric (Eq 5) formulation of Lanchester's equations without modifying Brown's analysis.

TABLE 14

PREDICTED AND OBSERVED BATTLE VICTORIES

R_f[a]	V_s[b]		A_s[c]	f_w[d]
	$y=0$	$y=1$		
Category I (31 Battles/Group)				
1.020	17.50	17.50	31	
1.071	13.14	13.19	17	.488
1.100	17.00	17.01	16	
1.128	19.00	19.01	19	
1.167	23.14	23.20	24	.343
1.200	19.00	19.00	22	
1.217	9.01	9.02	10	
1.250	16.01	16.02	18	.532
1.275	18.50	18.50	19	
1.320	16.99	16.95	15	
1.364	21.50	21.50	20	.371
1.400	19.99	20.00	22	
1.444	22.99	22.97	16	
1.500	22.08	22.12	17	.300
1.500	20.00	20.00	21	
1.563	18.00	18.00	20	
1.655	21.00	21.00	18	.344
1.667	22.00	22.00	23	
1.778	24.00	24.00	22	
1.833	23.00	23.00	22	.271
1.944	20.77	20.73	20	
2.000	21.50	21.50	18	
2.000	24.50	24.50	20	.269
2.250	22.00	22.00	18	
2.400	25.00	25.00	23	
2.667	24.50	24.50	19	.322
3.000	22.78	22.82	19	
3.409	23.00	23.00	20	
4.167	23.00	23.00	21	.269
6.714	22.00	22.00	19	
13–	8.00	8.00	6[e]	
Category II (13 Battles/Group)				
1.235	8.97	8.92	9	
1.571	7.00	7.00	7	
1.786	12.00	12.05	10	
2.143	12.00	12.00	12	
2.500	10.00	10.00	10	
3.000	12.00	12.00	11	
3.617	8.85	8.77	9	
5.000	9.50	9.51	10	
6.154	12.00	12.00	13	
7.750	12.00	12.09	13	
16.000	13.00	13.00	12	
71–	6.00	6.00	5[f]	

[a]Breakpoints in values of R_f used in forming groups.

[b]Victories to stronger predicted by Brown.

[c]Actual victories to stronger out of 31 battles per group in Category I and 13 in Category II.

[d]Frequency of victory to weaker predicted by Brown.

[e]Nine battles in this group.

[f]Six battles in this group.

surrounding army in the field. If all battles had been sorted by attacker and defender, a higher measure of accord might be found between fact and theory. The fact remains that the results reported here are completely independent of γ and only weakly dependent on force ratio.

7. RESUME

On account of the use made of Lanchester's equations in analyzing war games it would have been desirable to know who was the attacker in these battles and how many battles really ended in a draw. In spite of this omission it is believed that, by means of the foregoing tests, the following observations may be made for the battles studied:

(a) In general, force ratio has little to do with predicting the outcome of battle.

(b) Lanchester's square law is the poorest among poor alternative choices of deterministic laws—results are insensitive to γ.

(c) By elimination it is presumably the exchange ratio E that has controlled the outcome of battle and is in some fashion responsible for the accuracy with which Brown's form of the stochastic equations explains Table 3. Although the expression given by Eq 2 apparently yields by hindsight a satisfactory estimate of E, there is no theory for E.

The writer concludes that in the absence of any method of predicting E reliably there is little value in a simple version of Lanchester's equations as a predictive tool where the only known quantities are initial strengths.

APPENDIXES

Appendix A

PROOF OF CONVERGENCE OF SOLUTION OF STOCHASTIC EQUATIONS TO DETERMINISTIC FORM

For a square-law environment Morse and Kimball[10] give the following differential-difference equation for the probability $P(m, n, t)$ that at time t there remain m Red survivors and n Blue survivors

$$\frac{d}{dt} P(m, n, t) = mE^{1/2} \left[P(m, n+1, t) - P(m, n, t) \right] + nE^{-1/2} \left[P(m+1, n, t) - P(m, n, t) \right] \qquad \text{(A1)}$$

whose right-hand side can be rewritten:

$$= mE^{1/2} \frac{P(n+1) - P(n)}{(n+1) - n} + nE^{-1/2} \frac{P(m+1) - P(m)}{(m+1) - m}$$

$$= mE^{1/2} \frac{\Delta P}{\Delta n} + nE^{-1/2} \frac{\Delta P}{\Delta m}$$

in an obvious notation. If we now assume m and n to be continuous variables and allow Δm and Δn to approach zero, Eq A1 becomes

$$\frac{\partial}{\partial t} P(m, n, t) = mE^{1/2} \frac{\partial P}{\partial n} + nE^{-1/2} \frac{\partial P}{\partial m}. \qquad \text{(A2)}$$

Three elementary changes of variable bring about the further transformations:

New Variables		New Equation	
$x = mE^{1/2}$	$y = nE^{-1/2}$	$\frac{\partial P}{\partial t} = x \frac{\partial P}{\partial y} + y \frac{\partial P}{\partial x}$	(A3)
$u = x + y$	$v = x - y$	$\frac{\partial P}{\partial t} = u \frac{\partial P}{\partial u} - v \frac{\partial P}{\partial v}$	
$r = \log u$	$s = \log v$	$\frac{\partial P}{\partial t} = \frac{\partial P}{\partial r} - \frac{\partial P}{\partial s}$	(A4)

The canonic equation for Eq A4 is

$$\mathbf{a} \cdot \nabla P(r, y, z) = 0, \qquad \text{(A5)}$$

where $\mathbf{a}$ is any constant vector. Equation A5 says that the gradient of P vanishes in the direction of $\mathbf{a}$; or, P is a function of only the coordinates in the plane normal to $\boldsymbol{a}$. The projection of the vector $r = ix + jy + kz$ on this plane is

$$\boldsymbol{a} \times r = i(a_y z - a_z y) + j(a_z x - a_x z) + k(a_x y - a_y x),$$

and P must depend on x, y, and z only through those combinations that must appear in it. Since any of these components is a linear combination of the other two, there are only two degrees of freedom, and one may say that a general solution of Eq A5 is

$$P = P(a_y z - a_z y,\ a_x z - a_z x),$$

an arbitrary function of two of the coordinates of $\boldsymbol{a} \times r$.

Returning to Eq A4 and using the previous argument one finds that its solution can be written

$$P = P(t - s,\ t + r),$$

which with some manipulation becomes

$$P = P[e^{-t}(x - y),\ e^{t}(x + y)]. \quad \text{(A6)}$$

For a battle whose initiation at $t = 0$ is described by $x = x_0$ and $y = y_0$ the value of P then is $P(t = 0) = P(x_0 - y_0,\ x_0 + y_0)$. If the initial values of x and y are certain then

$$P(0) = 2 \cdot \delta[x - y - (x_0 - y_0)] \cdot \delta[x + y - (x_0 + y_0)] \quad \text{(A7)}$$

in conformity with Eq A6. The δ functions are the usual unit impulse functions in Dirac's notation.[18] It is easy to show that Eq A7 is more simply written

$$P(0) = \delta(x - x_0) \cdot \delta(y - y_0), \quad \text{(A8)}$$

which establishes the normalization given in Eq A7.

To get the complete solution that possesses Eqs A7 or A8 as initial value, substitute x_1 for $mE^{1/4}$ and y_1 for $nE^{-1/4}$ in Eq 1 and put γ and g equal to unity. The resulting differential equations

$$\frac{dx_1}{dt} = -y_1 \quad \text{and} \quad \frac{dy_1}{dt} = -x_1$$

have the solutions

$$\begin{Bmatrix} x_1 \\ y_1 \end{Bmatrix} = \begin{Bmatrix} x_0 \\ y_0 \end{Bmatrix} \cosh t - \begin{Bmatrix} y_0 \\ x_0 \end{Bmatrix} \sinh t$$

or

$$\left.\begin{aligned} e^{-t}(x_1 - y_1) &= x_0 - y_0 \\ e^{t}(x_1 + y_1) &= x_0 + y_0 \end{aligned}\right\} \quad \text{(A9)}$$

It is plausible (and provable) that the desired solution to Eq A3 is

$$P(x, y, t) = 2 \cdot \delta[e^{-t}(x - y) - (x_0 - y_0)] \cdot \delta[e^{t}(x + y) - (x_0 + y_0)],$$

which with the help of Eq A9 becomes

$$P(x, y, t) = 2 \cdot \delta[x - y - (x_1 - y_1)] \cdot \delta[x + y - (x_1 + y_1)]$$

or finally

$$P(x, y, t) = \delta(x - x_1) \cdot \delta(y - y_1) \tag{A10}$$

in the pattern of Eq A8. Utilizing the identity

$$\frac{d}{dx}[\delta(u)] = \frac{-\delta(u)}{u}\frac{du}{dx}$$

we get from Eq A10

$$\frac{\partial}{\partial t}P(x, y, t) = -\left(\frac{y_1}{x - x_1} + \frac{x_1}{y - y_1}\right)P(x, y, t)$$

$$y\frac{\partial}{\partial x}P = -\frac{y}{x - x_1}P = -\frac{y_1}{x - x_1}P$$

$$x\frac{\partial}{\partial y}P = -\frac{x}{y - y_1}P = -\frac{x_1}{y - y_1}P,$$

where the last expressions in the second and third lines follow from the known properties of δ functions. Thus Eq A10 satisfies Eq A3.

It has now been proved (a) that the solution to Eq A1 assigns zero probability to any values of m and n at any time t for which m and/or n is not given by the solution of Eq 1 with γ and g set equal to unity and (b) that by virtue of Eqs A6 and A9 it assigns the same measure of certainty to $P(m, n, t)$ that it does to $P(m_0, n_0, 0)$ provided m and n are solutions to Lanchester's square law evolving from initial values m_0 and n_0.

Finally it may be shown that the integral $\int_0^{\infty} P(x, y, t)\,dt$ is a δ function that vanishes unless $x = x_1$ and $y = y_1$ for every value of t, so that zero probability is assigned to any battle that departs at all from an ideal square-law behavior. QED

For values of γ different from 1, different methods must be used since Eq 1 cannot in general be integrated in closed form and no appropriate analog of Eq A9 exists.

Appendix B

CATEGORY I BATTLES THAT CONFORM MOST NEARLY TO LANCHESTER'S LINEAR OR SQUARE LAW

As indicated in the Summary and in Sec 5, a list of those battles of Category I that conform most nearly to Lanchester's linear or square law is presented here. (The spellings are Bodart's.)

Battle	Year	Winner	Loser
Wimpfen	1622	HRE,[a] Spain	German insurgents
Fleurus	1622	German insurgents	HRE, Spain
Mulhausen	1674	France	HRE
Lund	1676	Sweden	Denmark
St Denis	1678	France	Holland, Spain
Steenkerke	1692	France	Allies
Olaschin	1696	Turkey	HRE
Eeckeren	1703	France, Spain	England, Holland
Speyer	1703	France	HRE, Holland, Prussia
Hochstadt (Blenheim)	1704	England, Holland, Prussia, HRE	France, Bavaria
Turin	1706	HRE, Piedmont	France, Spain
Oudenarde	1708	England, Holland, HRE	France, Spain
Coni (Madonna dell'Olmo)	1744	France, Spain	Austria, Sardinia
Lawfeldt (Laffelt)	1747	France	England, Holland, Austria
Lobositz	1756	Prussia	Austria
Lutternberg	1758	France	Hesse, Hanover
Kunersdorf	1759	Russia, Austria	Prussia
Warburg	1760	England, Germany	France
Liegnitz	1760	Prussia	Austria
Torgau	1760	Prussia	Austria
Burkersdorf	1762	Prussia	Austria
Stillwater	1777	America	England
Neerwinden	1793	Austria	France
Wattignies	1793	France	Allies
Loano	1795	France	Austria, Sardinia
Neresheim	1796	France	Austria
Emmendiengen	1796	Austria	France
Caldiero	1796	Austria	France

[a] Holy Roman Empire.

Battle	Year	Winner	Loser
Verona[b]	1799		—
Magnano	1799	Austria	France
Novi	1799	Austria, Russia	France
San Martino d'Albaro	1800	France	Austria
Engen and Stockach	1800		
Monte Bocco	1800		
Marengo	1800		
Neuburg	1800		
Mincio	1800		
Gunzburg	1805		
Caldiero	1805	Austria	France
Preussisch-Eylau	1807	France	Russia, Prussia
Medina-de-Rio-Seco	1808	France	Spain
Espinosa-de-los-Monteros	1808	France	Spain
Sacile	1809	Austria	France
Regensburg	1809	France	Austria
Raab (Győr)	1809	France	Austria
Talavera de la Reina	1809	England, Spain	France
Schumla	1810	Turkey	Russia
Murviedro (Sagunto)	1811	France	Spain
Jakubovo	1812	Russia	France
Tschaschniki	1812	Russia	France
Wolkowisk (Izabelin)	1812	Austria, Saxony, France	Russia
Vitoria	1813	England, Portugal, Spain	France
Hagelberg	1813	Prussia	France
Bayonne	1813	England, Portugal, Spain	France
Brienne	1814	France	Prussia, Russia
Mincio	1914	France, Italy	Austria
Orthez	1814	England, Portugal	France
Lundy's Lane	1814	America	England
Wavre	1815	France	Prussia
Iganie	1831	Poland	Russia
Custozza	1848	Austria	Sardinia
Mortara	1849	Austria	Sardinia
Solferino	1859	France, Sardinia	Austria
Pea Ridge	1862	Union	Confederacy
Perryville	1862	Confederacy	Union
Prairie Grove	1862	Union	Confederacy
Murfreesboro (Stones River)	1862/3	Union	Confederacy
Ober-Selk and Jagel	1864	Austria	Denmark
Drewry's Bluff	1864	Confederacy	Union
Rossbrunn	1866	Prussia	Bavaria
Spicheren	1870	Germany	France
Bellevue	1870		
Amiens	1870		
Nuits	1870		
Schipka Pass	1877	Russia	Turkey
Liao-Jang	1904	Japan	Russia

[b]Indecisive battle between France and Austria. Austria assumed victorious.

REFERENCES

1. F. W. Lanchester, Aircraft in Warfare; The Dawn of the Fourth Arm, Constable & Co., Ltd., London, 1916.
2. H. K. Weiss, "Lanchester-Type Models of Warfare," in Proceedings of the First International Conference on Operational Research, Operations Research Society of America, Baltimore, 1957.
3. R. N. Snow, "Contributions to Lanchester Attrition Theory," RAND Corp., Project Rand, RA-15078, 1948.
4. R. L. Helmbold, "Lanchester Parameters for Some Battles of the Last Two Hundred Years (U)," Combat Operations Research Group, CORG-SP-122, 1961; "Historical Data and Lanchester's Theory of Combat," Combat Operations Research Group, CORG-SP-128, 1961.
5. Lewis F. Richardson, Statistics of Deadly Quarrels, Boxwood Press, Pittsburgh, 1960.
6. Quincy Wright, A Study of War, University of Chicago Press, Chicago, 1942.
7. Gaston Bodart (ed), Militär-historisches Kriegs-Lexikon, (1618–1905), C. W. Stern Vienna and Leipzig, 1908.
8. E. G. Smith and G. N. Donovan, "Trends in Warfare (U)," memo F.6, Army Operational Research Group, UK, 1955. CONFIDENTIAL—MODIFIED HANDLING
9. Randall E. Cline, "A Survey and Summary of Mathematical and Simulation Models as Applied to Weapons System Evaluation," Institute of Science and Technology, University of Michigan, Rept 3681-16-F, Ann Arbor, 1961.
10. P. M. Morse and G. E. Kimball, Methods of Operations Research, John Wiley & Sons, Inc., New York, 1950.
11. N. M. Smith, unpublished notes.
12. R. H. Brown, "A Stochastic Analysis of Lanchester's Theory of Combat," Operations Research Office (Research Analysis Corp.), ORO-T-323, Dec 55.
13. Harald Cramér, Mathematical Methods of Statistics, Princeton University Press, Princeton, N. J., 1946.
14. A. Hald, Statistical Theory with Engineering Applications, John Wiley & Sons, Inc., New York, 1952.
15. W. C. Suhler, "A Statistical Analysis of the Teichroew Pseudo-Random Normal Deviate Generator," Research Analysis Corp., RAC-TP-60, May 62.
16. P. A. M. Dirac, Quantum Mechanics, Oxford University Press, London, 1947, 3d ed